Lévy Flights, Non-local Search and Simulated Annealing

Ilya Pavlyukevich[*]

22 January 2007

Abstract

We solve a problem of non-convex stochastic optimisation with help of simulated annealing of Lévy flights of a variable stability index. The search of the ground state of an unknown potential is non-local due to big jumps of the Levy flights process. The convergence to the ground state is fast due to a polynomial decrease rate of the temperature.

Keywords: Lévy flights; simulated annealing; non-local search; heavy-tails; variable stability index; stable-like process; global optimisation.

1 Introduction

Let U be a potential function in $\mathbb{R}^d$ having several local minima and increasing fast at infinity. We look for a global minimum of U. Classical continuous-time simulated annealing (Boltzmann machine) (see [VG84, Gid, GH86, GM93]) consists in running a diffusion process

$$\hat{Z}_{0,z}(t) = z - \int_0^t \nabla U(\hat{Z}_{0,z}(u))\, du + \int_0^t \hat{\sigma}(u)dW(u),$$

$$\hat{\sigma}(t) = \left(\frac{\theta}{\ln(\lambda + u)}\right)^{1/2}, \tag{1.1}$$

where W is a standard Brownian motion, $\theta > 0$ denotes the cooling rate and $\lambda > 1$ parametrises the initial temperature, which equals $\sqrt{\theta/\ln(\lambda)}$ at time $t = 0$. It is known that there is a critical value $\hat{\theta}$ such that the diffusion $\hat{Z}(t)$ converges in distribution to the *global* minimum of U if $\theta > \hat{\theta}$, and the convergence fails otherwise. Moreover, the critical value $\hat{\theta}$ is the logarithmic rate of the principal non-zero eigenvalue $\lambda_1(\sigma)$ of a time homogeneous diffusion generator $A_\sigma f = \frac{\sigma^2}{2}\Delta f - \langle \nabla U, f \rangle$, i.e.

$$\hat{\theta} = -\lim_{\sigma \to 0} \sigma^2 \ln |\lambda_1(\sigma)|. \tag{1.2}$$

The value of $\hat{\theta}$ can be calculated explicitly, if one knows the heights of potential barriers between different wells of U (see [Wen72a, Wen72b] and [KM96] for precise results). Rigorous results on optimal cooling rate in simulated annealing algorithms can be found in [Haj88, CHS87, HKS89, HS90].

In order to accelerate the search, Szu and Hartley in [SH87] suggested the so-called *fast simulated annealing* (Cauchy machine), which is a combination of a classical Metropolis algorithm introduced in [MRRT53] and a concept of non-local search due to the heavy-tail Cauchy visiting distribution. The authors claimed that in the Cauchy machine the temperature can be chosen decreasing as a power of time, namely $\sigma(t) \sim t^{-1}$, and applied the algorithm in image processing, see [Szu93].

Motivated by [SH87], in our papers [Pav06b, Pav06a] we considered a continuous-time counterpart of the process $\hat{Z}$ driven by *Lévy flights* of stability index $\alpha \in (0, 2)$, and temperature $\sigma(t) \sim t^{-\theta}$, $\theta > 0$. We discovered that such a jump process never settles in the neighbourhood of a global minimum, but can be used to reveal a spatial structure of the potential U. The dynamics of Lévy flights with constant small noise was studied in our previous papers [IP06a, IP06b, IP06c].

---

[*]Correspondence should be addressed to pavljuke@mathematik-hu-berlin.de

In the present paper we solve the problem of global optimisation with help of *state-dependent* Lévy flights in a multi-well potential U. We show, that in certain annealing regimes, the global minimum is localised always, as in the classical Gaussian case. For simplicity, we restrict our theoretical argument in Sections 2–5 to one-dimensional potentials. However, it will be clear from the presentation, that the algorithm also works in a multi-dimensional setting. In our numerical example in Section 6 we consider a two-dimensional potential with five wells.

2   Results on the cooled down Lévy flights

In [Pav06b, Pav06a] we considered a one-dimensional Lévy flights process in an external potential U determined by the equation

$$Z_{0,z}^{(\alpha)}(t) = z - \int_0^t U'(Z_{0,z}^{(\alpha)}(u-))\,du + \int_0^t \frac{dL^{(\alpha)}(u)}{(\lambda+u)^\theta}. \tag{2.1}$$

We understand a Lévy flights process $L^{(\alpha)}$ as a symmetric stable Lévy process with stability index $\alpha \in (0,2)$, whose marginal distributions have the Fourier transform

$$\mathbf{E}e^{i\omega L^{(\alpha)}(t)} = e^{-c(\alpha)t|\omega|^\alpha} = \exp\left(t\int_{\mathbb{R}\backslash\{0\}} \left[e^{i\omega y} - 1 - i\omega y \mathbb{I}_D(y)\right] \frac{dy}{|y|^{1+\alpha}}\right), \tag{2.2}$$

where $\mathbb{I}_D(y)$ is the indicator function of the unit disk $D = \{y : |y| \leq 1\}$, and $c(\alpha) = 2|\cos\left(\frac{\pi\alpha}{2}\right)\Gamma(-\alpha)|$. We choose such a parametrisation in order to have a simple form of a Lévy measure $\nu(dy) = |y|^{-1-\alpha}dy$.

The measure ν is also called a jump measure of the process Lévy $L^{(\alpha)}$. It controls the intensity and sizes of its jumps. Indeed, let $\Delta L^{(\alpha)}(t) = L^{(\alpha)}(t) - L^{(\alpha)}(t-)$ denote the jump size of $L^{(\alpha)}$ at time instance $t > 0$. Then the number of jumps on the time interval $(0,t]$ with values in a set $J \subseteq \mathbb{R}$ is a Poissonian random variable with the mean $t\nu(J)$ (which can be possibly zero or infinite).

The process $L^{(\alpha)}$ is a Markov process with a non-local generator

$$Af(x) = \int_{\mathbb{R}\backslash\{0\}} \left[f(x+y) - f(x) - yf'(x)\mathbb{I}_D(y)\right] \frac{dy}{|y|^{1+\alpha}}, \tag{2.3}$$

which is also referred to as a fractional Laplacian, $A = \Delta^{\alpha/2}$.

We direct reader's attention to the books [Sat99, App04, Pro04] on a rigorous mathematical theory of Lévy processes and stochastic differential equations. Physical results on the subject can be found in [MK04, CGK$^+$04, BS02].

We assume that the potential U has n local minima m_i and $n-1$ local maxima s_i such that $-\infty = s_0 < m_1 < s_1 < \cdots < m_n < s_{n+1} = +\infty$. The extrema are non-degenerate, i.e. $U''(m_i) > 0$ and $U''(s_i) < 0$. Moreover we demand that $|U'(x)| > |x|^{1+c}$ as $|x| \to +\infty$ for some positive c.

In the small temperature limit, i.e. when $\lambda \to +\infty$ or $t \to +\infty$, the process $Z^{(\alpha)}$ can be seen as a random perturbation of a deterministic dynamical system

$$X_x^0(t) = x - \int_0^t U'(X_x^0(s))\,ds. \tag{2.4}$$

We denote $\Omega_i = (s_{i-1}, s_i)$, $1 \leq i \leq n$, the domains of attraction of the stable points m_i.

The positive parameter θ is called the cooling rate, and $\lambda > 0$ determines the initial temperature of the system which equals to $\lambda^{-\theta}$ at $t = 0$.

Equation (2.1) describes a non-linear dynamics of a Lévy particle, whose temperature is being decreased at a polynomial rate as $t \to \infty$.

In [Pav06b, Pav06a], we discovered two cooling regimes — slow cooling $\theta < 1/\alpha$ and fast cooling $\theta > 1/\alpha$ — in which the transitions of a particle between the wells of U have different asymptotic properties.

Let $\Delta > 0$ be a small number, and let $B_i = \{y : |y - m_i| \leq \Delta\}$ denote a Δ-neighbourhood of a local minimum m_i. Consider transition times

$$T_{s,z}^i = \inf\{u \geq s : Z_{s,z}^{(\alpha)}(u) \in \cup_{j\neq i}B_j\} \tag{2.5}$$

between different neighbourhoods B_i and the corresponding transition probabilities $\mathbf{P}_{s,z}(Z^{(\alpha)}(T^{i,\lambda}) \in B_j)$, $i \neq j$. Then for $\theta < 1/\alpha$ and $z \in B_i$ we have

$$\frac{\mathbf{E}_{0,z}T^i}{\lambda^{\alpha\theta}} \to [q_i^{(\alpha)}]^{-1}, \quad \lambda \to +\infty, \tag{2.6}$$

$$\mathbf{P}_{0,z}(Z^{(\alpha)}(T^i) \in B_j) \to q_{ij}^{(\alpha)}[q_i^{(\alpha)}]^{-1}, \quad i \neq j, \tag{2.7}$$

where

$$q_{ij}^{(\alpha)} = \int_{\Omega_j} \frac{dy}{|m_i - y|^{1+\alpha}} = \frac{1}{\alpha}\left|\frac{1}{|s_{j-1} - m_i|^\alpha} - \frac{1}{|s_j - m_i|^\alpha}\right|, \quad i \neq j,$$
$$q_i^{(\alpha)} = \int_{\mathbb{R}\setminus\Omega_i} \frac{dy}{|m_i - y|^{1+\alpha}} = \sum_{j\neq i} q_{ij}^{(\alpha)} = \frac{1}{\alpha}\left(\frac{1}{|s_{i-1} - m_i|^\alpha} + \frac{1}{|s_i - m_i|^\alpha}\right). \tag{2.8}$$

We have also shown that in the limit $t \to +\infty$, $Z_{0,z}^{(\alpha)}(t)$ has a distribution

$$\pi^{(\alpha)}(dy) = \sum_{i=1}^{n} \pi_i^{(\alpha)} \delta_{m_i}(dy) \tag{2.9}$$

where the vector $\pi^{(\alpha)} = (\pi_1^{(\alpha)}, \ldots, \pi_n^{(\alpha)})^T$, solves the equation $Q^T\pi^{(\alpha)} = 0$, $Q = (q_{ij}^{(\alpha)})_{i,j=1}^n$, $q_{ii}^{(\alpha)} = -q_i^{(\alpha)}$. It is clearly seen, that all $\pi_i^{(\alpha)} > 0$, and $Z^{(\alpha)}$ does not settle down near the global minimum of U. However, the values $\pi_i^{(\alpha)}$, which can be estimated from the Monte Carlo simulations, reveal the spatial structure of U, e.g. the sizes of the domains Ω_i.

If the cooling rate θ is above the threshold $1/\alpha$, the Lévy particle $Z^{(\alpha)}$ gets trapped in one of the wells and thus the convergence fails. Consider the first exit time from the i-th well

$$S_{s,z}^i = \inf\{u \geq s : Z_{s,z}^{(\alpha)} \notin \Omega_i\}. \tag{2.10}$$

Then, for $z \in B_i$, $1 \leq i \leq n$,

$$\mathbf{P}_{0,z}(S^i < \infty) = \mathcal{O}\left(\frac{1}{\lambda^{\alpha\theta-1}}\right), \quad \lambda \to \infty, \tag{2.11}$$

and consequently, $\mathbf{E}_{0,z}S^i = \infty$.

3 Lévy flights with variable stability index (stable-like processes)

In order to take into account the energy geometry of the potential, we have to make the Lévy flights process depend on its current position. Thus instead of Lévy flights $L^{(\alpha)}$ defined in (2.2), we consider now the so-called *stable-like process* $H = (H(t))_{t\geq 0}$, which is a Markov process defined by the non-local generator

$$Bf(x) = \int_{\mathbb{R}\setminus\{0\}} [f(x+y) - f(x) - yf'(x)\mathbb{I}_D(y)]\frac{dy}{|y|^{1+\alpha(x)}}, \tag{3.1}$$

with a function $\alpha(x)$ taking values in the interval $(0,2)$. Sometimes, the notation $B = \Delta^{\alpha(\cdot)/2}$ is used. The main difference between $L^{(\alpha)}$ and H consists in a dependence of a stable-like jump measure $\nu_x(dy) = |y|^{-1-\alpha(x)}$ on the spatial coordinate x. Thus, if $H(t_0) = x_0$, the instant jump distribution of H at time t_0 is governed by a stable measure $\nu_{x_0}(dy)$.

Of course, the dynamics of H is completely determined by a variable stability index $\alpha(x)$. For example, if $\alpha(x) = \alpha_0 \in (0,2)$, then H is just a usual Lévy flights process of index α_0. From now on, we assume that $\alpha(x)$ takes values strictly between 0 and 2, i.e. $0 < a \leq \alpha(x) \leq A < 2$, to exclude degeneration of the jump measure.

We are going to study the dynamics of a stochastic differential equation with the driving process H, namely

$$Y_{0,y}(t) = y - \int_0^t U'(Y_{0,y}(u-))\,du + \int_0^t \frac{dH(Y_{0,y}(u-),u)}{(\lambda+u)^\theta}. \tag{3.2}$$

3

For a better understanding of the process Y, it is instructive to consider a discrete time analogue of (3.2) given by the recurrent formula

$$y_k = y_{k-1} - U'(y_{k-1})h + \frac{\xi_k^h(y_{k-1})}{(\lambda + (k-1)h)^\theta}, \quad k \geq 1. \tag{3.3}$$

The discrete time dynamical system (3.3) is obtained from the Euler approximation of (3.2) with the time step h, and can be used for simulations. (However, one should be careful when U' is not globally Lipschits.) The random input is determined by the random variables $\xi_k(y)$ such that

$$\xi_k^h(y) \stackrel{d}{=} L^{(\alpha(y))}(h) \stackrel{d}{=} h^{1/\alpha(y)} L^{(\alpha(y))}(1) \tag{3.4}$$

where $L^{(\alpha(y))}(1)$ has a standard symmetric $\alpha(y)$-stable distribution with the Fourier transform (2.2).

4 One-well dynamics. Transitions and trapping

The dynamics Y of the Lévy flights with variable stability index in the force field U' is a result of an interplay of two independent effects. First, for small temperatures, i.e. when $\lambda \to +\infty$ or $t \to +\infty$, Y is close to the underlying deterministic trajectory X^0. Starting from any point of Ω_i, it follows X^0 with the same initial point and with high probability reaches a small neighbourhood of m_i in relatively short time. On the other hand, Y tries to deviate from X^0 making jumps controlled by the jump measure $\nu_x(dy)$. Finally, if Y is in the well Ω_i, it spends most of the time in a neighbourhood of m_i and thus has jumps approximately governed by the stable jump measure $\nu_{m_i}(dy) = |y|^{-1-\alpha(m_i)}dy$.

Thus, the exit time and the exit probability from the well Ω_i of the process Y are approximately the same as for the process $Z^{(\alpha(m_i))}$. This resemblance becomes exact if we consider a piece-wise constant stability index $\alpha(x)$, $\alpha(x) = \sum_{i=1}^n \alpha_i \mathbb{I}\{x \in \Omega_i\}$, $0 < \alpha_i < 2$. With this choice of $\alpha(x)$, the process Y is just driven by the equation (2.1) until it exits the well. (We omit a discussion on the behaviour of the process in the small neighbourhoods of the saddle points.)

Let us introduce the following transition and exit times for the process Y:

$$\tau_{s,z}^i = \inf\{u \geq s : Y_{s,z}(u) \in \cup_{j \neq i} B_j\}, \tag{4.1}$$

$$\sigma_{s,z}^i = \inf\{u \geq s : Y_{s,z} \notin \Omega_i\}. \tag{4.2}$$

It follows from (2.6) and (2.7), that if $y \in B_i$ and $\alpha(m_i)\theta < 1$, then the following relations hold as $\lambda \to +\infty$:

$$\frac{\mathbf{E}_{0,y}\tau^i}{\lambda^{\alpha(m_i)\theta}} \to [q_i^{(\alpha(m_i))}]^{-1}, \tag{4.3}$$

$$\mathbf{P}_{0,z}(Y(T^i) \in B_j) \to q_{ij}^{(\alpha(m_i))}[q_i^{(\alpha(m_i))}]^{-1}, \quad i \neq j. \tag{4.4}$$

On the other hand, if $\alpha(m_i)\theta > 1$, the Lévy particle gets trapped in the well due to (2.11), i.e.

$$\mathbf{P}_{0,z}(\sigma^i < \infty) = \mathcal{O}\left(\frac{1}{\lambda^{\alpha(m_i)\theta - 1}}\right), \quad \lambda \to +\infty, \tag{4.5}$$

and consequently, $\mathbf{E}_{0,z}\sigma^{i,\lambda} = \infty$.

Relations (4.3), (4.4) and (4.5) are crucial for our analysis.

5 Non-local random search and simulated annealing

5.1 Getting trapped in an assigned well

We demonstrate now how to drive the Lévy particle Y to an assigned well Ω_i, if the approximate location of its minimum m_i is known.

Indeed, the function $\alpha(x)$ given, the global dynamics of Y is determined by the values $\alpha(m_i)$, $1 \leq i \leq n$, and the cooling rate θ. Moreover, for our analysis we can freely choose both $\alpha(x)$ and θ.

Let $\alpha(x)$ be smooth and attain its unique global maximum at m_i. Then we have

$$\alpha(m_i) > \max_{j \neq i} \alpha(m_j). \tag{5.1}$$

For instance, one can take $\alpha(x) = (A - a)/(1 + (x - m_i)^2) + a$ for some $0 < a < A < 2$. Then we can choose $\theta > 0$, such that

$$\alpha(m_i)\theta > 1, \quad \text{whereas} \quad \alpha(m_j)\theta < 1 \quad \text{for} \quad j \neq i. \tag{5.2}$$

With this choice of parameters, as $t \to +\infty$, the particle leaves any well Ω_j, $j \neq i$, in a finite time according to (4.3). Moreover, since all transition probabilities in (4.4) are strictly positive, the probability to enter the well Ω_i after a finite number of transitions between the wells Ω_j, $j \neq i$, equals 1. Finally, upon entering Ω_i, the particle gets trapped there due to (4.5).

5.2   Looking for the global minimum

Let M be the (unique) unknown global minimum of the potential U. To make a Lévy particle settle near it, we have to determine the appropriate variable stability index $\alpha(x)$ and the cooling rate θ such that

$$\alpha(M)\theta > 1, \quad \text{whereas} \quad \alpha(m_i)\theta < 1, \ m_i \neq M. \tag{5.3}$$

Let $\varphi(u)$ be an *arbitrary* smooth monotone *decreasing* function on $\mathbb{R}$, $0 < a < \varphi(u) < A < 2$. Then we set

$$\alpha(y) = \varphi(U(y)). \tag{5.4}$$

It is clear that

$$\alpha(M) > \alpha(m_i) \quad \text{for all} \quad m_i \neq M, \tag{5.5}$$

and we choose θ to satisfy relations (5.3). The trapping of the particle near the global minimum M follows from the argument of the preceding section.

5.3   The local minimum with maximal energy

Analogously, we can determine the coordinate of the (unique) 'highest' local minimum m, i.e. such that

$$U(m) = \max_{1 \leq i \leq n} U(m_i). \tag{5.6}$$

In this case, we should take $\alpha(y) = \psi(U(y))$ with an *arbitrary* smooth monotone *increasing* function ψ, $0 < a < \varphi(y) < A < 2$, which leads to the inequality

$$\alpha(m) > \alpha(m_i) \quad \text{for all} \quad m_i \neq m, \tag{5.7}$$

and the arguments of the previous sections justify the success of the search.

5.4   A local minimum with certain energy

Finally, we can perform a search not for the global minimum of U but for a local minimum m_E satisfying the condition $U(m_E) \leq E$ for some energy level E. In such a case we take a piece-wise constant stability index

$$\alpha(y) = \begin{cases} A, & U(y) \leq E, \\ a, & U(y) \geq E, \quad 0 < a < A < 2, \end{cases} \tag{5.8}$$

and a cooling rate θ satisfying conditions $A\theta > 1$ and $a\theta < 1$. Consequently, if for some i, the minimum $U(m_i)$ lies below the threshold E, then near this minimum Y behaves like a jump-diffusion $Z^{(A)}$ driven by a Lévy flights process $L^{(A)}$, and thus the Lévy particle gets trapped due to (4.5). If the well minimum lies above the level E, a Lévy particle behaves like a process $Z^{(a)}$ and leaves this well in finite time due to relations (4.3) and (4.4).

If there are several wells with minima below E, the Lévy particle settles down near one of them.

Running the search for decreasing energies $E_1 > E_2 > ...$, we can also estimate the ground energy level $E^* = U(M)$ and to determine the global minimum M. Analogously, one can determine a local minimum m_E satisfying conditions $U(m_E) \geq E$ or $E_1 \leq U(m_E) \leq E_2$.

We emphasise that the search algorithm described in this section requires no particular information about the potential U: to determine the local minimum m_E we use three numbers a, A and θ, the energy level E, and values of $U'(x)$.

6 Numerical example

In the preceding sections we gave a theoretic justification of the search algorithm in dimension one. It is clear that a d-dimensional case, $d > 1$, does not differ much. It suffices to consider isotropic (spherically symmetric) d-dimensional Lévy flights with a jumping measure $\nu_{\mathbf{x}}(d\mathbf{y}) = |\mathbf{y}|^{-d-\alpha(\mathbf{x})} d\mathbf{y}$, and to calculate new values of transition probabilities in (2.8) writing d instead of 1 in the integrals.

Similarly to (3.3), we generate a d-dimensional Markov chain

$$\mathbf{y}_k = \mathbf{y}_{k-1} - \nabla U(\mathbf{y}_{k-1})h + \frac{\xi_k^h(\mathbf{y}_{k-1})}{(\lambda + (k-1)h)^\theta}, \quad 1 \leq k \leq n, \tag{6.1}$$

with some initial value $\mathbf{y}_0$. Isotropic α-stable random vectors $\xi_k^h(\mathbf{y})$ are obtained as marginals of a subordinated standard d-dimensional Brownian motion (e.g., see Examples 24.12 and 30.6 in [Sat99]). More precisely, let S be an $\alpha/2$-stable strictly positive random variable with the Fourier transform

$$\mathbf{E} \exp(i\omega S) = \exp\left(-|\omega|^{\alpha/2}(1 - i\,\mathrm{sgn}(\omega)\tan(\frac{\pi\alpha}{4}))\right), \quad \omega \in \mathbb{R}, \tag{6.2}$$

and the Laplace transform

$$\mathbf{E} \exp(-uS) = \exp\left(-\frac{u^{\alpha/2}}{\cos(\frac{\pi\alpha}{4})}\right), \quad u \geq 0. \tag{6.3}$$

Let $\mathbf{W}$ be a standard Gaussian vector independent of S with a characteristic function

$$\mathbf{E} \exp\left(i\langle\omega, \mathbf{W}\rangle\right) = \exp\left(-\frac{|\omega|^2}{2}\right), \quad \omega \in \mathbb{R}^d. \tag{6.4}$$

Then one obtains a standard isotropic d-dimensional α-stable random vector $\mathbf{L}^{(\alpha)}$ with the Fourier transform

$$\mathbf{E} \exp\left(i\langle\omega, \mathbf{L}^{(\alpha)}\rangle\right) = \exp\left(\int_{\mathbb{R}^d\setminus\{0\}} \left[e^{i(\omega,\mathbf{y})} - 1 - i(\omega,\mathbf{y})\mathbb{I}_D(\mathbf{y})\right] \frac{d\mathbf{y}}{|\mathbf{y}|^{d+\alpha}}\right)$$
$$= \exp\left(-\frac{\pi^{d/2}}{2^\alpha} \frac{|\Gamma(-\frac{\alpha}{2})|}{\Gamma(\frac{d+\alpha}{2})}|\omega|^\alpha\right), \quad \omega \in \mathbb{R}^d, \tag{6.5}$$

as

$$\mathbf{L}^{(\alpha)} \stackrel{d}{=} c(\alpha, d)\sqrt{S} \cdot \mathbf{W}, \quad c(\alpha, d) = \frac{1}{\sqrt{2}}\left[\pi^{d/2}\cos\left(\frac{\pi\alpha}{4}\right)\frac{|\Gamma(-\frac{\alpha}{2})|}{\Gamma(\frac{d+\alpha}{2})}\right]^{1/\alpha} \tag{6.6}$$

Finally the random increments in (6.1) can be calculated as

$$\xi_k^h(\mathbf{y}) = h^{1/\alpha(\mathbf{y})}\mathbf{L}^{(\alpha(\mathbf{y}))}. \tag{6.7}$$

6.1 Five-well potential in $\mathbb{R}^2$

To illustrate the efficiency of the method, we consider a two-dimensional potential function U given by the formula

$$U(\mathbf{y}) = \left[1 - \frac{1}{1 + 0.05(y_1^2 + (y_2 - 10)^2)} - \frac{1}{1 + 0.05((y_1 - 10)^2 + y_2^2)} - \frac{1.5}{1 + 0.03((y_1 + 10)^2 + y_2^2)}\right.$$
$$\left. - \frac{2}{1 + 0.05((y_1 - 5)^2 + (y_2 + 10)^2)} - \frac{1}{1 + 0.1((y_1 + 5)^2 + (y_2 + 10)^2)}\right](1 + 0.0001(y_1^2 + y_2^2)^{1.2}). \tag{6.8}$$
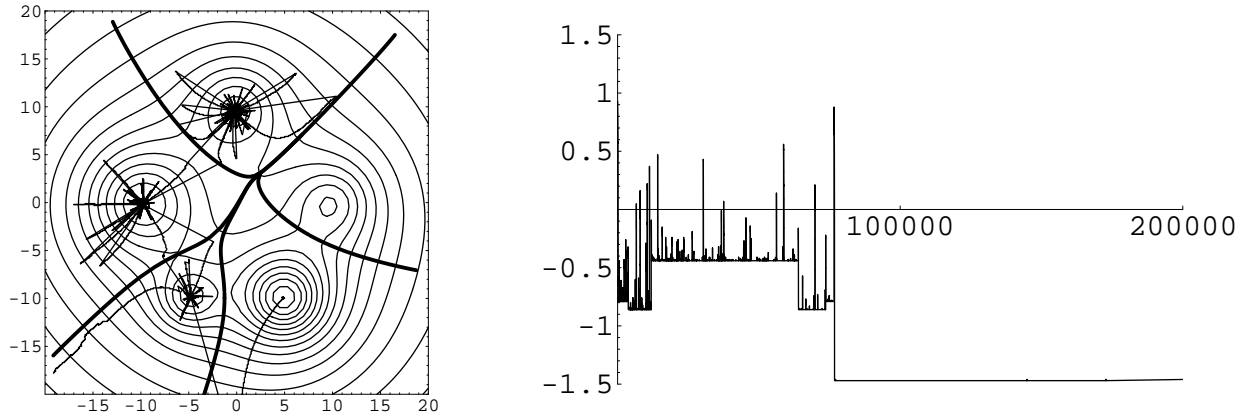
Figure 1: A typical random search path $(\mathbf{y}_k)_{0 \le k \le n}$ of a Lévy particle (l.) and the values of the potential function $U(\mathbf{y}_k)$ (r.). Thick lines on the left figure denote the boundaries of the attraction domains Ω_i of the stable points $\mathbf{m}_i$, $i = 1, \ldots, 5$.

The function U has five local minima $\mathbf{m}_i$, $i = 1, \ldots, 5$, with the following coordinates and energy values:

$$
\begin{aligned}
\mathbf{m}_1 &\approx (-9.73, -0.11), & U(\mathbf{m}_1) &\approx -0.85, \\
\mathbf{m}_2 &\approx (-0.09, \ \ 9.64), & U(\mathbf{m}_2) &\approx -0.44, \\
\mathbf{m}_3 &\approx (\ \ 9.59, -0.37), & U(\mathbf{m}_3) &\approx -0.54, \\
\mathbf{m}_4 &\approx (\ \ 4.92, -9.89), & U(\mathbf{m}_4) &\approx -1.46, \\
\mathbf{m}_5 &\approx (-4.79, -9.79), & U(\mathbf{m}_5) &\approx -0.79.
\end{aligned}
\tag{6.9}
$$

We perform a search for a local minimum $\mathbf{m}_E$ having the energy less than $E = -1$, i.e. such that $U(\mathbf{m}_E) \le -1$. In our example, $\mathbf{m}_E = \mathbf{m}_4$. According to Section 5.4, we choose a piece-wise constant stability index

$$
\alpha(\mathbf{y}) = \begin{cases} 1.8, & \text{if } U(\mathbf{y}) < -1, \\ 1.1, & \text{if } U(\mathbf{y}) \ge -1, \end{cases}
\tag{6.10}
$$

and $\theta = 0.75$.

We perform 100 simulations of the Markov chain (6.1) for $n = 2 \cdot 10^6$, initial conditions $\mathbf{y}_0$ distributed uniformly in the square $[-20, 20]^2$, the initial temperature $\lambda = 10^4$, and the time step $h = 0.1$. The global minimum $\mathbf{m}_4$ was determined in 96 cases ($U(\mathbf{y}_n) = -1.46$). The local minimum $\mathbf{m}_3$ was located twice, and each of the minima $\mathbf{m}_1$ and $\mathbf{m}_2$ once. A typical random path $(\mathbf{y}_k)_{0 \le k \le n}$ on the plain and the corresponding values of the energy function $U(\mathbf{y}_k)$ are shown on Figure 1.

We emphasise, that in our simulations we used only values of the potential U without any additional information about its geometry.

7 Conclusion and discussion

In this paper we presented a new stochastic algorithm for global optimisation. It allows to determine a global minimum of an unknown potential U with help of simulated annealing of non-Gaussian jump-diffusions driven by the so-called stable-like processes, or Lévy flights with a variable stability index $\alpha(x)$. We have shown that choosing $\alpha(x)$ in an appropriate way, we can force the Lévy particle to settle in a neighbourhood of the global maximum of U. We note, that the non-constant behaviour of the stability index is crucial for the success of the search, and a similar algorithm with usual spatially homogeneous Lévy flights, i.e. when $\alpha(x) = \alpha_0$, leads to quite different results, see [Pav06a].

Our method has the following advantages in comparison with the Gaussian simulated annealing considered in the introduction. First, the search of the global minimum is non-local, i.e. when the annealed process

7

leaves a potential well, it does not necessarily pass to one of the neighbouring wells, but with strictly positive probability can jump to any well. Moreover, the probability to jump into the deepest well is maximal, if this well is also spatially the biggest, which is observed in typical potential landscapes, see [Sch97]. We do not expect that our algorithm would effectively detect the so-called 'golf-hole' wells. Mean transition times between the wells increase as a power of the large parameter λ or, equivalently, the current time t. We can easily obtain theoretic estimates for a number of transitions between the wells before settling in the deepest well. These estimates follow from the analysis of a discrete time Markov chain with transition probabilities $p_{ij} = q_{ij}^{(\alpha(m_i))}[q_i^{(\alpha(m_i))}]^{-1}$, $p_{ii} = 0$.

Second, we have more freedom to choose the parameters of the system. Indeed, if the values of $U(m_i)$ are not known, there is no method which helps to determine the cooling rate θ. (One has the same problem to determine $\hat{\theta}$ in (1.2) in Gaussian case.) However, in our algorithm, θ is chosen together with a variable stability index $\alpha(x)$.

Third, our method allows to drive the Lévy particle into any well whose location is approximately known. We can also determine a local minimum with energy below the certain given value. The choice of parameters in these regimes is independent of the geometry of the potential. Such search regimes are not possible in the classical setting.

Finally, the temperature decreases polynomially fast in time, i.e. $\sim t^{-\theta}$, and not logarithmic. This significantly increases the accuracy of empirical estimates for the local minima locations.

Although the theoretic basis for the success of the search is established, many questions are still open. For example, we have to understand how to choose the optimal pair $\alpha(x)$ and θ, which minimises the search time. Indeed, if m_i is not a global minimum, we can reduce the life time of the particle in the neighbourhood of m_i making $\alpha(m_i)$ small. On the other hand, in this case, the process Y will tend to make very big jumps, and thus jump out to one of the peripheral wells. As a consequence, the search would be slow, if the global minimum of U is attained in one of the inner wells. Thus, the value of $\alpha(m_i)$ should not be very small to exclude very big jumps, and should be well separated from $\alpha(M)$ to block trapping in the false well. The problem of very big jumps can be also avoided by consideration of truncated Lévy flights with maximal jump size not exceeding the size of the search domain. However, in this case the simulation of random input can be more complicated. We shall address these and other questions in our further research.

References

[App04]   D. Applebaum. *Lévy Processes and Stochastic Calculus*, volume 93 of *Cambridge Studies in Sdvanced Mathematics*. Cambridge University Press, Cambridge, 2004.

[BS02]   D. Brockmann and I. M. Sokolov. Lévy flights in external force fields: from models to equations. *Chemical Physics*, 284:409–421, 2002.

[CGK+04] A. V. Chechkin, V. Yu. Gonchar, J. Klafter, R. Metzler, and L. V. Tanatarov. Lévy flights in a steep potential well. *Journal of Statistical Physics*, 115(5–6):1505–1535, 2004.

[CHS87]   T.-S. Chiang, C.-R. Hwang, and S.-J. Sheu. Diffusion for global optimization in $\mathbb{R}^n$. *SIAM Journal on Control and Optimization*, 25(3):737–753, 1987.

[GH86]   S. Geman and C.-R. Hwang. Diffusions for global optimization. *SIAM Journal on Control and Optimization*, 24(5):1031–1043, 1986.

[Gid]   B. Gidas. Global optimization via the Langevin equation. In *Conference on Decision and Control, 24th, Fort Lauderdale, FL, USA, 11–13 Dec. 1985.*

[GM93]   S. B. Gelfand and S. K. Mitter. Metropolis-type annealing algorithms for global optimization in $\mathbb{R}^d$. *SIAM Journal on Control and Optimization*, 31(1):111–131, 1993.

[Haj88]   B. Hajek. Cooling schedules for optimal annealing. *Mathematics of Operations Research*, 13(2):311–329, 1988.

[HKS89]   R. Holley, S. Kusuoka, and D. Stroock. Asymptotics of the spectral gap with applications to the theory of simulated annealing. *Journal of Functional Analysis*, 83(2):333–347, 1989.

[HS90]   C.-R. Hwang and S.-J. Sheu. Large-time behavior of perturbed diffusion Markov processes with applications to the second eigenvalue problem for Fokker-Planck operators and simulated annealing. *Acta Applicandae Mathematicae*, 19(3):253–295, 1990.

[IP06a]   P. Imkeller and I. Pavlyukevich. First exit times of SDEs driven by stable Lévy processes. *Stochastic Processes and their Applications*, 116(4):611–642, 2006.

[IP06b]   P. Imkeller and I. Pavlyukevich. Lévy flights: transitions and meta-stability. *Journal of Physics A: Mathematical and General*, 39:L237–L246, 2006.

[IP06c]   P. Imkeller and I. Pavlyukevich. Metastable behaviour of small noise Lévy-driven diffusions. *ArXiv Mathematics e-prints*, math.PR/0601771, 2006. Submitted.

[KM96]   V. N. Kolokol'tsov and K. A. Makarov. Asymptotic spectral analysis of a small diffusion operator and the life times of the corresponding diffusion process. *Russian Journal of Mathematical Physics*, 4(3):341–360, 1996.

[MK04]   R. Metzler and J. Klafter. The restaurant at the end of the random walk: recent developments in the descroption of anomalous transport by fractional dynamics. *Journal of Physics A: Mathematical and General*, 37:R161–R208, 2004.

[MRRT53] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, and E. Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(5):1087–1092, 1953.

[Pav06a] I Pavlyukevich. Cooling down Lévy flights. *ArXiv e-prints*, cond-mat/0701651, 2006. Submitted.

[Pav06b] I Pavlyukevich. Simulated annealing of Lévy-driven jump-diffusions, 2006. Submitted.

[Pro04]   Ph. E. Protter. *Stochastic integration and differential equations*, volume 21 of *Applications of Mathematics*. Springer, Berlin, second edition, 2004.

[Sat99]   K.-I. Sato. *Lévy processes and infinitely divisible distributions*, volume 68 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999.

[Sch97]   J. C. Schön. Preferential trapping on energy landscapes in regions containing deep-lying minima: The reason for the success of simulated annealing? *Journal of Physics A: Mathematical and General*, 30(7):2367–2389, 1997.

[SH87]   H. Szu and R. Hartley. Fast simulated annealing. *Physics Letters A*, 122(3,4):157–162, 1987.

[Szu93]   H. Szu. Automated fault recognition by image correlation neutral network technique. *IEEE Transactions on Industrial Electronics*, 40(2):197–208, 1993.

[VG84]   D. Vanderbilt and Louie S. G. A Monte Carlo simulated annealing approach to optimization over continuous variables. *Journal of Computational Physics*, 56:259–271, 1984.

[Wen72a] A. D. Wentzell. On the asymptotic behavior of the greatest eigenvalue of a second-order elliptic differential operator with a small parameter in the higher derivatives. *Soviet Math doklady*, 13(1):13–17, 1972.

[Wen72b] A. D. Wentzell. On the asymptotic of eigenvalues of matrices with elements of order $\exp\{-V_{ij}/(2\varepsilon^2)\}$. *Soviet Math doklady*, 13(1):65–68, 1972.

Ilya Pavlyukevich
Department of Mathematics
Humboldt University of Berlin
Rudower Chaussee 25
12489 Berlin Germany
E.mail: pavljuke@mathematik.hu-berlin.de
http://www.mathematik.hu-berlin.de/~pavljuke